

СТВОРЕННЯ ПРОГРАМНОГО ПАКЕТУ ДЛЯ ПРЕДСТАВЛЕННЯ СТАТИСТИЧНИХ ДАНИХ

М.В. КОЛІСНИЧЕНКО^{1*}, Г.Ю. СИДОРЕНКО²

^{1.} *магістрант кафедри САІТ, НТУ «ХПІ», Харків, УКРАЇНА*

^{2.} *доцент кафедри САІТ, канд. техн. наук, НТУ «ХПІ», Харків, УКРАЇНА*

** email: marina.kolesnichenko6732@gmail.com*

Розвиток складних комп'ютерних програм вплинуло на підйом нових методів прогнозування і аналізу. В сучасних умовах, коли інформаційні потоки стали особливо масивними, з'являється колосальна кількість даних. Також, з огляду на те, що швидкість всіх процесів в суспільстві зростає, збільшується і потреба в виявленні швидких відповідей на виникаючі питання, з метою пошуку яких і проводиться величезна кількість досліджень і збір даних в усіх сферах життя. Зараз ринок статистичного програмного забезпечення вражає своїм різноманіттям, незважаючи на його специфічність, існує більш тисячі різноманітних програм вирішують завдання статистичного аналізу даних.

Метою даної роботи є аналіз пакетів статистики, а також розробка програмного забезпечення для роботи зі статистичними даними.

Всі програми статистичної обробки даних за ознакою функціональності можна розділити на професійні, універсальні та спеціалізовані. Статистичні програми відносяться до наукомісткого програмного забезпечення, ціна їх часто недоступна індивідуальному користувачеві. Професійні пакети мають велику кількість методів аналізу, універсальні пакети – кількість функцій, достатню для універсального застосування. Спеціалізовані ж пакети орієнтовані на якусь вузьку область аналізу даних.

Всі статистичні програмні пакети мають свої переваги та недоліки. Вони створюються та використовуються безпосередньо для задач статистики, прогнозування тощо. Наприклад, пакет IBM SPSS Statistics є одним з кращих в світі програмних рішень для статистичного аналізу. Він містить інструменти прогнозного моделювання і просунуті аналітичні функції для вирішення комерційних і дослідницьких завдань. IBM SPSS Statistics охоплює всі стадії аналітичного процесу, від підготовки даних і управління ними до виконання аналізу і створення звітів. Це програмне рішення надає персоналізовану функціональність і настроюються інтерфейси для різних рівнів кваліфікації та функціональних обов'язків користувачів [1].

Користувачі SPSS Statistics отримують доступ до великої кількості статистичних функцій, що дає їм можливість проводити складні види аналізу навіть при відсутності навичок володіння складною мовою програмування.

Цей пакет є фактичним стандартом для багатьох підприємств, дослідних інститутів і фахівців за статистикою. До недоліків пакету SPSS варто віднести:

1. Висока ціна в порівнянні зі статистичними пакетами також рівня.

2. Відсутність реалізації деяких важливих тестів часових рядів (зокрема - тестів на стаціонарність).

3. Маленький набір математичних функцій перетворення змінних.

Статистична середа R – є безумовним лідером серед некомерційних систем статистичного аналізу, а також при проведенні науково-технічних розрахунків в більшості західних університетських центрів і багатьох провідних фірмах [2].

Переваги пакета R:

1. Існують реалізації пакету під операційні системи сімейств Microsoft Windows, Mac OS X, Linux.

2. Базова комплектація R займає небагато місця і містить всі функції, необхідні для проведення статистичного аналізу.

3. Хороша графічна візуалізація подання даних та результатів їхнього аналізу.

Недоліки мови R:

1. Труднощі навчання в програмі через ручне введення великої кількості команд.

2. Другий недолік R - відносна повільність. Деякі функції, особливо ті які, використовують цикли, і види об'єктів, а саме списки і таблиці даних, працюють в десятки разів повільніше, ніж їх аналоги в комерційних пакетах.

Тому беручи до уваги вище описані переваги та недоліки постає задача для розробки власного пакету для обробки та візуалізації статистичних даних, яке буде містити базу даних, що призначена для зберігання інформації, моніторинг отриманих параметрів, формування звіту та аналіз мінливості отриманих параметрів.

Статистичний пакет також повинен задовольняти наступному набору вимог:

1. Використання простої проблемно-орієнтованої мови для формулювання завдання користувача.

2. Автоматична організація процесу обробки даних.

3. Ведення банку даних користувача і складання звіту про результати проведеного аналізу.

4. Діалоговий режим роботи користувача з пакетом.

5. Сумісність з іншим програмним забезпеченням

В роботі проведено дослідження технологічного процесу створення програмного пакету для візуалізації статистичних досліджень. Також необхідна розробка системи прийняття рішень, що буде створюватися на аналізі формування звітності, що відображає стан технологічного процесу.

Список літератури:

1. Aldrich J. Using IBM SPSS Statistics: An Interactive Hands-On Approach/ J. Aldrich // Los Angeles: SAGE – 2015.

2. Argyrous G. Statistics for Research: With a Guide to SPSS/G. Argyrous// SAGE Publications Ltd. – 2011.